基于概念格理论的产品领域本体构建研究*

陆佳莹[1] 袁勤俭[1] 黄 奇[2] 钱韵洁[1]
[1](南京大学信息管理学院 南京 210093)
[2](南京大学国家信息资源管理南京研究基地 南京 210093)

摘要:【目的】基于概念格理论探索并提出复用产品分类的产品领域本体构建方法，对以往复用产品分类的产品领域本体构建方法进行改进。【方法】从产品分类中抽取类、属性及关系，并对其结构进行合理调整，基于概念格理论生成形式背景及概念格，并以此为基础自动生成产品领域本体。【结果】本文方法消除了产品分类本身存在的层级不分明的局限性，可以解决产品分类定义的属性间的冗余问题，提升本体的可重用性和可共享性，并通过可视化的形式呈现了现有和潜在的实体与关系。【局限】仅选取医疗产品构建的一个片段来说明该方法的具体实施方案，并仅探讨该方法对于 eCl@ss 产品分类体系的适用性。【结论】基于概念格构建产品领域本体前需对本体的领域和范围进行明确界定，依据科学分类原则对类目进行梳理与归类，在进行本体构建时构建属性词典并定义对象属性，通过形式背景与概念格的生成自动生成本体。

关键词: 概念格 领域本体 本体构建 产品本体
分类号: TP391 G35

1 引 言

概念格(Concept Lattices)，又称伽罗瓦格(Galois Lattices)由 Wille 于 1982 年首次提出[1]，它是形式概念分析中的核心数据单元。在哲学层面上，概念被理解为由外延和内涵两个基本单元组成。在这一思想的基础之上，概念格将哲学中的概念形式化，认为概念格的每一个节点都是一个形式概念，它包含两个组成部分：外延，即概念所包含的所有实例的集合；内涵是概念的描述，即概念所覆盖实例所共有特征(属性)的集合[2]。此外，概念格是一种由偏序集诱导的格结构，显示出了概念间的泛化和例化关系，而形式概念连同它们之间的泛化和例化关系共同构成了一个概念格。一个概念格的图形由称为哈斯图(Hasse Diagram)的线图表示，哈斯图是一个偏序集的图形表示，是通过向上的箭头来反映偏序集的包含关系。通过哈斯图能够生动展现概念格中概念与概念间的关系，从而实现对数据的可视化。总而言之，概念格模型是数据分析与知识发现的有力工具，目前已被广泛应用于本体工程、知识发现、数字图书馆、语义检索、人工智能、软件开发[3]等领域。

2 相关研究

早期的本体构建方法通常是基于专家的知识结构手工构建本体，一方面手工构建本体成本高、耗时费力、可移植性差，另一方面基于专家经验确立概念层次时很可能造成概念的冗余[4]。而生成概念格的过程本质上是一种基于实例–属性的概念聚类过程，能够通过映射自动生成清晰的本体概念层次，并以此作为本体构建与生成的基础，在概念层次上对本体形式化表示方法进行补充与完善，避免概念的冗余。因此，基于概念格构建的本体具备良好的语法与完备的语义，能够支持基于领域文本的智能化推理，从而提高本体的可共享性和可重用性。

鉴于概念格在处理概念化数据方面独特的优势，越来越多的学者将概念格理论应用于领域本体构建当

通讯作者：陆佳莹，ORCID：0000-0002-4534-6967，E-mail：njulujy@163.com。
*本文系国家社会科学基金重点项目“基于知识组织的产品分类本体研究”(项目编号:13ATQ005)的研究成果之一。

中，这些研究主要可以分成两类：

(1) 从非结构化数据生成概念格的领域本体构建方法。这类方法通常是用自然语言处理的方法对非结构化的文本信息(如：网页信息、文本等)进行处理，从中提取概念层次生成概念格，最后基于概念格自动生成本体。Haav[5]提出一种基于概念格的本体构建方法，这种方法主要是通过自然语言的方法从内容比较短的领域文本中提取形式背景，将形式概念分析中的概念和本体中的概念直接等同从而得到概念格，并通过形式概念方法和概念格缩减从形式背景计算得到初始本体。与 Haav 的方法不同，Cimiano 等[6]提出的方法则将形式概念分析中的属性与本体的概念相等同，使用自然语言解析的方法得到形式背景并构建概念格，最后由概念格生成领域本体。上述两种方法主要采用领域内的短文本作为本体学习的数据源，刘萍等[7]选取 CNKI 全文数据库的期刊文章作为本体学习的数据源，结合关联规则挖掘和形式概念分析的方法来识别概念之间的相关关系，以此作为情报学本体构建的基础。

(2) 从结构化数据生成概念格的领域本体构建方法。这类领域本体构建方法主要依赖主题词表、叙词表、分类法、关系数据库等结构化数据，Nanda 等[8]提出产品元件的本体设计方法 PFODM，该方法从材料明细表中提取概念与属性，并利用形式概念分析生成概念层次，最后基于概念层次生成概念格并生成领域本体。而欧阳纯萍等[9]提出的本体构建方法则依赖关系数据库，获得由属性和领域概念构成的形式背景并生成概念格，最后自动生成领域本体。此外，滕广青等[10]以结构化资源交通汉语主题词表为基础，并结合相关领域文本，通过并置运算得到异构资源的概念格，最后自动生成交通领域本体。

通过文献调研可以发现，学者将概念格理论引入领域本体的构建中，进行了一系列有价值的研究。这些方法利用概念格辅助获取结构化的数据信息，从已知的结构化或非结构化数据集中提取有用的概念层次，并以此作为构建本体的基础。近年来学者对于从非结构化数据生成概念格的本体构建方法进行了广泛而深入的研究，提出一系列基于自然语言处理和机器学习技术的概念层次提取方法，而对于从结构化数据生成概念格的领域本体构建方法则主要面向关系数据库以及传统图书馆资源组织方式(如：叙词表、分类法、主题词表等)。对于产品领域本体的构建而言，关系数据库以及传统图书馆资源组织方式显然是不适用的，而产品分类这类结构化资源能为产品领域本体的构建提供指导[11]。然而，过去在产品分类的指导下进行本体的概念层次生成时主要依赖专家经验，鲜有将概念格理论引入产品分类的研究。

因此，本研究选取概念格理论作为研究的理论基础，更加深入地探讨如何复用产品分类构建产品本体，并试图提出普适性更强的产品领域本体构建方法，促使商家对产品和服务信息进行有效的组织和管理，实现消费者与商家之间信息与知识的共享。

3　基于概念格理论的产品领域本体构建流程

结合产品领域本体的特征，本文借鉴七步法[12]和骨架法[13]的思想，设计出产品领域本体的构建流程，如图 1 所示：

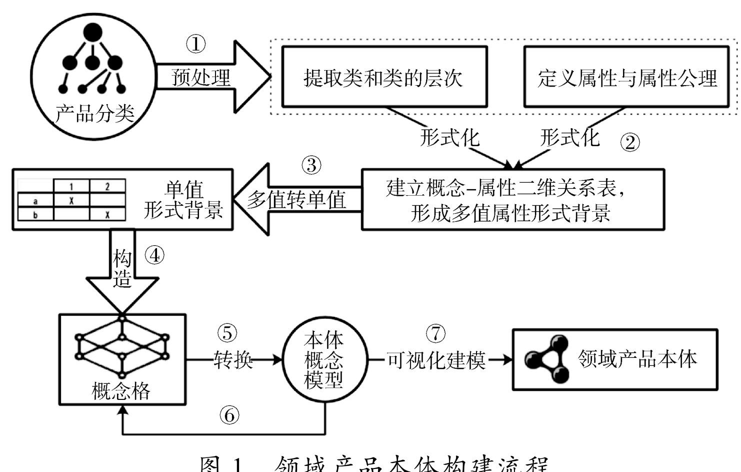


图 1　领域产品本体构建流程

主要概括为：产品分类预处理、形式背景生成与概念格构造和本体原型生成及形式化描述三个阶段，具体遵循以下 7 个步骤:

(1) 对产品分类进行预处理，从中提取类和类的层次，明确产品类目之间的种属、部分与整体、组合关系等，并将产品属性转化为本体中的数据属性和对象属性，通过对象属性界定类与属性以及属性间的关系;

(2) 将预先定义好的类、属性及关系形式化为概念–属性二维关系表，并形成多值形式背景，一个多值形式背景可以用四元组(G, M, W, I)表示，它由对象集合 G、多值属性集合 M、属性值域 W 以及这几个集合之间的一个三元关系 $I(I \in G \times M \times W)$ 构成;

(3) 由于形式概念分析方法处理的对象是单值形式背景，需要将多值形式背景转化为单值形式背景，而单值形式背景可以用一个三元组(G, M, I)表示，它由对象集合 G、单值属性集合 M 以及一个二元关系 $I(I \in G \times M)$ 构成;

(4) 利用美国国家癌症研究所编写的 OWLFCAView Protégé 插件[14]可以自动将单值形式背景转换为构建本体所需要的概念格;

(5) OWLFCAView Protégé 插件能够将概念格转换为本体的概念模型，由这个概念模型能够得到一个本体原型;

(6) 由本体设计师或领域专家对本体概念模型进行迭代确认与修正，直到最终生成的本体原型能够满足需求为止;

(7) 利用 Protégé 工具对最终确认的本体概念模型进行可视化建模，最终生成领域产品本体。

4 基于概念格的产品领域本体构建实例

4.1 产品分类预处理

选用 eCl@ss 产品分类作为本体构建的数据源。eCl@ss 是德国创建产品和服务的分类标准，是目前世界较为先进的且使用较为广泛的一套适用于电子商务环境的产品分类标准[15]。eCl@ss 产品与服务分类体系共有 30 个大类，涵盖了约 38 000 种产品分类和 16 000 种产品标签，基本上囊括了目前所有在使用的产品和服务。具体而言，本文选取 eCl@ss 产品分类 8.1 版本中的第 34 大类“医药和医疗技术”作为本体构建复用的产品分类体系，通过定义类和类的层次以及定义属性与属性公理对该产品分类进行预处理。

(1) 定义类和类的层次

eCl@ss 分类体系在“医药和医疗技术”这一类别下总共分为 52 个类目，该大类在 eCl@ss 中的体系结构如图 2 所示:

34 Medicine, medical technology
- 34-01 Medicine (pharmacopeia)
- 34-05 Active substance for pharmaceutical
- 34-06 Active substance for veterinary drugs
- 34-07 Active substance for plant protection
- 34-08 Active substance for cosmetics
- 34-14 Glove (medical)
- 34-15 Clothing (medical device, patient protection)
- 34-18 Plaster, medical tape
- 34-19 Retention, compression
- 34-20 Immobilization therapy supplies
- 34-21 Wound care
- 34-22 Injection, infusion or transfusion supplies
- 34-23 Angiographic supplies, therapeutical vascular catheter
- 34-24 Anesthesia, intensive care, dialysis or haemofiltration supplies
- 34-25 Urologic supplies
- 34-26 Radiologic supplies
- 34-27 Functional diagnostics supplies
- 34-28 Endoscopy, endotherapy, minimally invasive surgery (MIS)
- 34-29 Drainage, biopsy and puncture equipment
- 34-30 Medical instrument
- 34-31 Surgical equipment (apparatus)
- 34-32 Implant
- 34-33 Wound closure supplies
- 34-34 Medical equipment (instruments and small medical devices)
- 34-35 Surgical covers and drapes
- 34-36 Individual set
- 34-37 Ostomy care
- 34-38 Incontinence care
- 34-39 Patient positioning supplies
- 34-40 Patient care supplies
- 34-41 Sterilization equipment
- 34-42 Enteral technique
- 34-43 Physical rehabilitation devices
- 34-44 Physiotherapeutic equipment
- 34-45 Hospital furniture
- 34-46 Alternative medicine article
- 34-47 Medical aid
- 34-48 Hemotherapy
- 34-49 Light/microscope (medical)
- 34-50 Vital functions/Intensive care medicine device
- 34-51 Mass transfer device (medical)
- 34-52 Functional diagnosis device
- 34-53 Patient monitoring equipment
- 34-54 Therapy device
- 34-55 Diagnostic imaging system
- 34-56 Radiotherapy/tumor therapy
- 34-57 X-ray radiation (surveillance)
- 34-58 Surgical device system
- 34-59 Endoscope, endoscopy equipment (medical)
- 34-60 Medical specialist equipment
- 34-61 In-hospital transportation
- 34-62 Clinical Cleaning Device (medical)

图 2 医药和医疗技术大类二级子类体系结构

eCl@ss 与其他产品分类体系相比，具有描述精度高、通用性强等优势，同时也具有类目间的层次关系局限于四级类目下而不够科学合理的劣势。因此，在进行二级类目梳理时将不受 eCl@ss 四级类目的限制，而依据科学分类的原则进行类目划分，同时保留 eCl@ss 原有的分类编号，确保构建的产品本体的通用性和兼容性。

本文对于 eCl@ss 中34大类——医药和医疗技术类目中所有类目进行了调整，总体调整策略如图3所示。

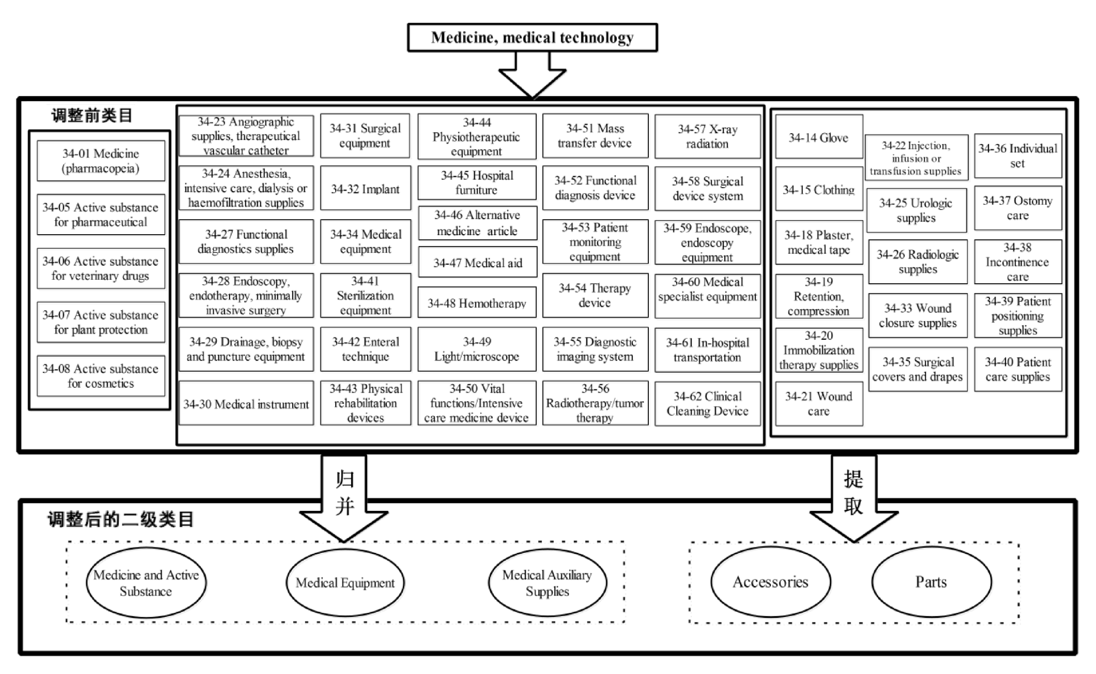


图 3　二级类目整体调整方法示意图

具体而言，新建五个二级类目包括：药品与活性物质(Medicine and Active Substance)、医疗器械(Medical Equipment)、医疗辅助用品(Medical Auxiliary Supplies)、部件(Parts)和配件(Accessories)。然后，将 34 大类下的 52 个二级类目分别归并至药品与活性物质类、医疗器械和医疗辅助用品类中，并从中原二级类目下的"parts"和"accessories"下提取类目分别放置部件类和配件类中。

(2) 定义属性与属性公理

在 OWL 中属性主要分为数据属性(Data Property)和对象属性(Object Property)，其中数据属性主要用来描述对象与具体取值之间的关系，而对象属性则主要用来描述对象与对象之间的关系[16]。

eCl@ss 通过属性和值对产品和服务信息进行描述，列举了实体拥有的基础属性(BSP)和特别属性(SSP)，并为属性提供了定义。而从本体构建的层面上看，eCl@ss 给出的产品和服务的属性全部都是数据属性，因此在构建本体时可以复用这类数据属性，而不需要重新定义。以 34-47"医疗辅助器材"下的 34-47-01-01"隐形眼镜"为例，eCl@ss 给出了该四级类目的制造商名称、产品名称、产品风险等级、所含材料的重量等属性，在构建医疗产品本体时，可以将这些属性直接复用为 OWL 中的数据属性。在定义数据属性时，还需要对数据类型进行定义，目前 Protégé 中支持将数据类型定义为布尔类型(boolean)、浮点类型(float)、整型类型(int)和字符串类型(string)。同样以 34-47"医疗辅助器材"下的 34-47-01-01"隐形眼镜"为例，在定义制造商名称、产品名称这类数据属性时，需要选择 string 类型，定义所含材料重量这类数据属性时，需要选择 float 数据类型，而对于产品风险等级则选择 int 数据类型。

而对象属性主要反映对象与对象之间的关系，eCl@ss 中涉及到的关系定义包括：类与属性的关系、部分与整体关系、配件与器件的关系。因此，本研究定义了三种对象属性如表 1 所示：

表 1　对象属性定义及其内涵

关系	对象属性定义	内涵
类与属性关系	has_property	表示某个类拥有某个属性
部分与整体关系	is_part_of	表示某个类是其他类的部件
配件与器件关系	is_accessory_of	表示某个类是其他类的配件

其中，has_property 表示类与属性的关系，例如："制造商名称"、"产品名称"等均是"隐形眼镜"的属性，它们之间的关系可以用对象属性"has_property"表示；is_part_of 表示部分与整体间的关系，该类目表示的均是某个类目是其他类目的部件，例如："检查镜"是"小

型医疗设备”的一部分，它们之间的关系可以用对象属性 is_part_of 表示；is_accessory_of 表示配件与器件之间的关系，配件与部件的不同之处在于配件不一定为器件所独有，它有一定的通用性，例如：“压舌板”、“耳注射器”等均是“小型医疗设备”的配件，它们之间的关系可以用对象属性 is_accessory_of 表示。

4.2 生成形式背景及构造概念格

eCl@ss 中的产品和服务属性存在着交叉部分，因此为了更为清晰地说明类与属性间的关系，避免冗余，本研究将属性、部件和配件分别建立二级类目，编写表示存在关系的规则语句“has_property some values from Property(or Parts or Accessories)”调用属性、部件和配件，这条规则语句表示存在量词约束，表示某个类目至少存在一个或者某些特定的属性来自“Property (or Parts or Accessories)”大类。在编写完整语义规则的基础之上，使用 OWLFCAView Protégé 插件，自动生成形式背景，并构造概念格。

本研究在生成形式背景并构造概念格时，可分为三种情形：

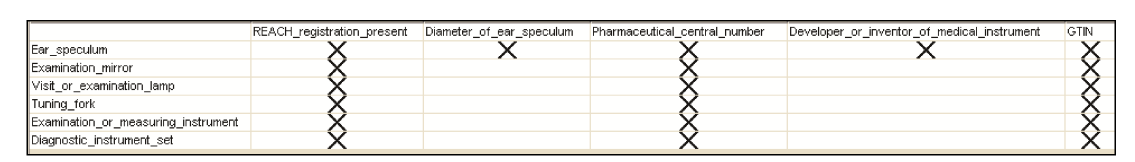
①基于产品类–属性的关系生成形式背景并构造概念格。这种情形下 eCl@ss 中定义的产品类目对应形式背景中的对象，四级产品类目下的属性对应形式背景中的属性，而关系指的是某个产品类拥有某个属性；

②基于产品类–部件类的关系生成形式背景并构造概念格。这种情形下产品类目对应形式背景中的对象，形式背景中的属性对应部件类目，而关系是指某个产品类拥有某个部件；

③基于产品类–配件类生成形式背景并构造概念格。这种情形下产品类目对应对象，配件类目对应属性，而关系指某个产品类拥有某个配件。

(1) 基于产品类–属性的形式背景及概念格

首先构建属性词典，将“医疗器械”类下的所有属性都以“Property”类的形式存储；编写规则语句“has_property some values from Property”，表示某个类目至少存在一个或者某些特定的属性来自“Property”大类，能够将每一个产品类和与其相对应的属性类进行关联；通过 OWLFCAView Protégé 插件可以提取 34-34 类“医疗器械(工具和小型医疗器械)”下某个子类的形式背景，本文选取“医疗器械(工具和小型医疗器械)”类目的子类 34-34-04 类“医疗器械工具”构造如图 4 所示的形式背景，其中行表示类，列表示属性，“×”表示某个类拥有某个属性(如：“检查镜”这个类目拥有“REACH 注册编号”、“治疗中心编号”和“全球贸易项目代码(GTIN)”三个属性)，并在此基础之上生成相应的概念格，如图 5 所示。



	REACH_registration_present	Diameter_of_ear_speculum	Pharmaceutical_central_number	Developer_or_inventor_of_medical_instrument	GTIN
Ear_speculum	×	×	×	×	×
Examination_mirror	×		×		×
Visit_or_examination_lamp	×		×		×
Tuning_fork	×		×		×
Examination_or_measuring_instrument	×		×		×
Diagnostic_instrument_set	×		×		×

图 4　产品类–属性形式背景图(部分)

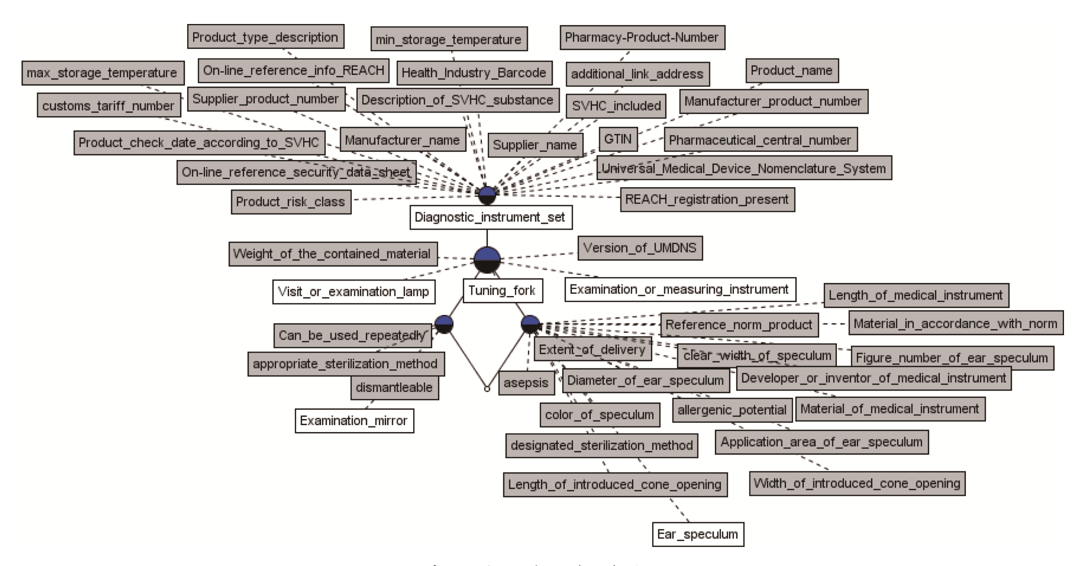
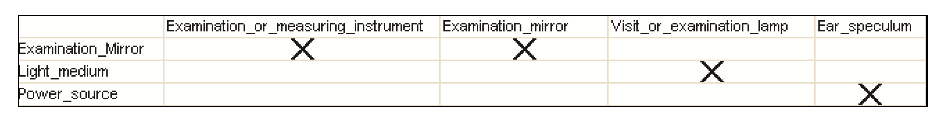


图 5　产品类–属性概念格示例图

(2) 基于部件类的形式背景及概念格

在 34-34“医疗器械”类目下，部件有“电源”、“检查镜”和“光媒介”三种。编写规则语句“has_parts some values from Parts”，它表示某个类目至少存在一个或者某些特定的部件来自“Parts”大类，能够将“检查镜”、“光介质”和“电源”这三种部件与“工具类”下的类目进行关联，最后生成如图 6 所示的部件类形式背景图，其中行表示类，列表示部件，“×”表示某个类拥有某个部件(如：“耳镜”这个类目拥有“电源”这个部件)，并在此基础上生成如图 7 所示的概念格。



	Examination_or_measuring_instrument	Examination_mirror	Visit_or_examination_lamp	Ear_speculum
Examination_Mirror	X	X		
Light_medium			X	
Power_source				X

图 6　部件类形式背景图

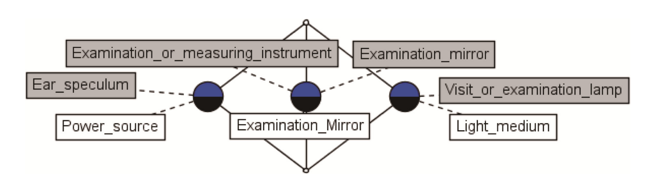


图 7　部件类概念格示例图

(3) 基于配件类的形式背景及概念格

与部件不同，配件类不仅要求一个类是另一个类的部分，而且要求作为部分的类必须被其他类所共有，即作为配件的类必须具有一定的通用性。依据这个原则，提取“医疗器械”类目下的全部配件，并与产品类进行组配。编写规则语句“has_accessories some values from Accessories”，表示某个类目至少存在一个或者某些特定的部件来自“Accessories”大类，能够将“耳镜”、“舌压板”等部件与“工具类”下的类目进行关联，生成如图 8 所示的部件类形式背景图，其中行表示类，列表示配件，“×”表示某个类拥有某个配件(如：“检查和测量工具”这个类目拥有“舌压板”这个配件)，并在此基础上生成如图 9 所示的概念格。

	Ear_Speculum	Diagnostic_instrument_set	Examination_mirror	Examination_or_measuring_instrument
Ear_syringe	X			
Ear_Speculum	X			
Case_suitcase		X		
Air_douche_otoscope_insufflator	X			
Tongue_swab		X	X	
Mirror_warmer				X
Dry_powder_insufflator		X		
Barany_noise_box		X		
Tongue_depressor		X	X	

图 8　配件类形式背景图

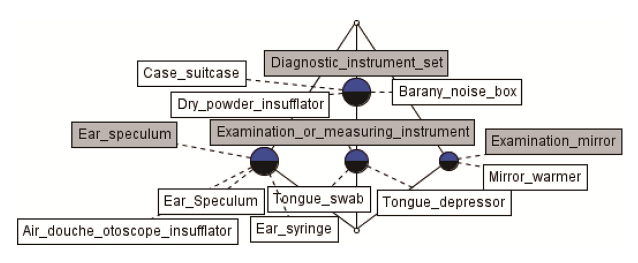


图 9　配件类概念格示例图

4.3　本体原型生成及形式化描述

OWLFCAView Protégé 插件能够将概念格转换为本体的概念模型，由这个概念模型能够得到一个本体原型。由领域专家对本体概念模型进行迭代确认与修正，通过 Protégé 生成最终的医疗产品本体和相应的 OWL DL 文档，截取形式化后的医疗产品 OWL DL 文档编码的片段，如下所示。该 OWL DL 编码片段表示“舌压板(Tongue Depressor)”是配件类(Accessories)的子类，并且“舌压板”是“检查与测量工具(Examination or Measuring Instrument)”和“诊断工具套件(Diagnostic Instrument Set)”的配件。

```
<?xml version="1.0"?>
<rdf:RDF
    xmlns:xsp="http://www.owl-ontologies.com/2005/08/07/
      xsp.owl#"
    xmlns:swrlb="http://www.w3.org/2003/11/swrlb#"
    xmlns:swrl="http://www.w3.org/2003/11/swrl#"
    xmlns:protege="http://protege.stanford.edu/plugins/owl/
      protege#"
    xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
    xmlns:xsd="http://www.w3.org/2001/XMLSchema#"
    xmlns="http://www.owl-ontologies.com/Ontology
      1430730323.owl#"
    xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#"
    xmlns:owl="http://www.w3.org/2002/07/owl#"
  xml:base="http://www.owl-ontologies.com/Ontology
    1430730323.owl">
  <owl:Ontology rdf:about=""/>
  <owl:Class rdf:ID="Tongue_ depressor">
    <rdfs:subClassOf>
      <owl:Restriction>
        <owl:someValuesFrom>
          <owl:Class rdf:ID="Examination_or_measuring_
            instrument"/>
        </owl:someValuesFrom>
        <owl:onProperty>
          <owl:ObjectProperty rdf:ID="is_accessory_of"/>
        </owl:onProperty>
      </owl:Restriction>
    </rdfs:subClassOf>
    <rdfs:subClassOf>
      <owl:Restriction>
```

```
        <owl:someValuesFrom>
          <owl:Class rdf:ID="Diagnostic_instrument_set"/>
        </owl:someValuesFrom>
        <owl:onProperty rdf:resource="#is_accessory_of"/>
      </owl:Restriction>
    </rdfs:subClassOf>
    <rdfs:subClassOf>
      <owl:Class rdf:ID="Accessories"/>
    </rdfs:subClassOf>
  </owl:Class>
```

5 结　语

本研究以概念格理论为基础，探讨如何复用产品分类eCl@ss构建产品本体，提出一种基于产品分类的普适性更强的产品本体构建方法，这与以往学者基于构建流程经验上认知的本体构建方法有很大的区别。研究提出的方法基于成熟的理论框架，将概念格理论引入产品本体构建过程中。并且，该方法能够对产品和服务信息进行形式化的、通用的知识表示，从而满足企业进行信息管理的需求。除此之外，研究对提出的基于概念格的产品本体构建方法选取本体构建的实例进行说明，探索并提出了建立该类本体的技术实现方案，为本体工程师构建同类型的产品本体提供参考。本研究得到的结论与启示如下:

(1) 在构建本体前对本体的领域和范围进行明确界定能够从根源上减少本体构建的冗余，提高构建本体的可共享性和可复用性。现存的本体在领域和范围方面存在大量交叉与重复。为了解决这个问题，有学者指出可以通过本体间的映射对相同主题的本体进行语义互联。而本研究发现要从根源上解决本体构建中的冗余问题，需要在本体构建前对本体的领域和范围进行明确界定，查找是否已有相同领域和范围的本体。

(2) 依据科学分类原则对类目进行梳理与归类，能够消除产品分类存在的层级不分明的局限性。以本文选用的产品分类eCl@ss为例，与其他产品分类体系相比，eCl@ss具有描述精度高、通用性强等优势，同时也具有类目间的层次关系局限于四级类目下而不够科学合理的劣势，因此在构建本体前对类目进行梳理是有必要的。其他产品分类也具有类似的特点，需要本体工程师依据科学分类的原则对所有类目进行梳理和归类，使得构建的本体依据科学分类的原则具有较好的通用性。

(3) 在进行本体构建时构建属性词典并定义对象属性，能够解决产品分类定义的属性间的冗余问题，清晰地说明类与属性间的关系。eCl@ss产品分类在四层分类的底层对产品和服务的属性进行描述，这些属性之间存在交集，若直接将这些属性复用至本体中则会造成冗余。此外，eCl@ss给出了产品和服务类目的数据属性，但并未给出描述类目间关系的对象属性。因此，在构建本体时本体工程师可以将所有eCl@ss定义的数据属性定义成属性词典，在Protégé中放入专门属性类目下，并通过定义对象属性关联产品类与属性类，从而更清晰地说明产品类与属性间的关系。

(4) 通过形式背景与概念格的生成，能够提升本体的可重用性和可共享性，并通过可视化的形式呈现现有和潜在的实体与关系。传统基于自然语言的本体虽易于构建，但缺少形式化的语义。而本研究发现基于概念格的本体构建能够提供形式化的语义，实现本体间的重用与共享。一方面，形式背景由对象、属性和关系的三元组构成，能够为本体提供形式化的信息组织模式；另一方面，概念格是基于形式背景生成的哈斯图，能够通过可视化的形式呈现现有和潜在的实体与关系。除此之外，本研究在构建产品本体的实例时发现，在复用产品分类生成形式背景和概念格时，通常可以分为三种情况：基于产品类–属性的关系，这种情形下产品类目对应形式背景中的对象，而产品类目下的属性对应形式背景中的属性，而关系指的是某个产品类拥有某个属性；基于产品类–部件类的关系，这种情形下产品类目对应形式背景中的对象，形式背景中的属性对应部件类目，而关系是指某个产品类拥有某个部件；基于产品类–配件类，这种情形下产品类目对应对象，配件类目对应属性，而关系指某个产品类拥有某个配件。

未来的研究还可以从以下三个方面进行:

(1) 选取不同的领域对该方法进行应用，从而进一步验证该方法在实际应用环境中的可用性;

(2) 基于概念格理论提出产品本体映射的相关方法，对异构的产品本体进行语义互联;

(3) 可以围绕复用其他标准产品分类(例如：UNSPSC、GPC等)的本体构建方法展开，进一步提高研究广度与深度。

参考文献：

[1] Wille R. Restructuring Lattice Theory: An Approach Based on Hierarchies of Concepts [A]. // Ivan Rival. Ordered Sets

[M]. Netherlands: Springer, 1982: 445-470.
[2] 谢志鹏，刘宗田．概念格与关联规则发现[J]．计算机研究与发展，2000，37(12)：1415-1421．(Xie Zhipeng, Liu Zongtian. Concept Lattice and Association Rule Discovery [J]. Journal of Computer Research and Development, 2000, 37(12): 1415-1421.)
[3] 毕强，滕广青．国外形式概念分析与概念格理论应用研究的前沿进展及热点分析[J]．现代图书情报技术，2010 (11): 17-23. (Bi Qiang, Teng Guangqing. Analysis of the Progress and Hotspots in Applied Research of FCA and Concept Lattice Theory Abroad [J]. New Technology of Library and Information Service, 2010 (11): 17-23.)
[4] 胡可云，陆玉昌．概念格及其应用进展[J]．清华大学学报：自然科学版，2000，40(9)：77-81．(Hu Keyun, Lu Yuchang. The Concept Lattice and Its Application Progress [J]. Journal of Tsinghua University: Science and Technology, 2000, 40(9): 77-81.)
[5] Haav H M. A Semi-Automatic Method to Ontology Design by Using FCA [C]. In: Proceedings of the 2nd International Workshop on Concept Lattices and their Application, 2004: 13-24.
[6] Cimiano P, Hotho A, Stumme G, et al. Conceptual Knowledge Processing with Formal Concept Analysis and Ontologies [C]. In: Proceedings of the 17th International Conference on Industrial and Engineering Application of Artificial Intelligence and Expert System. 2004: 189-207.
[7] 刘萍，胡月红．基于FCA和关联规则的情报学本体构建[J]．现代图书情报技术，2012(2): 34-40. (Liu Ping, Hu Yuehong. Development of Domain Ontology in Information Science Based on FCA and Association Rules [J]. New Technology of Library and Information Service, 2012(2): 34-40.)
[8] Nanda J, Simpson T W, Kumara S R, et al. A Methodology for Product Family Ontology Development Using Formal Concept Analysis and Web Ontology Language [J]. Journal of Computing and Information Science in Engineering, 2006, 6(2): 103-113.
[9] 欧阳纯萍，胡长军，李扬，等．一种基于 FCA 的面向关系数据库的本体学习方法[J]．计算机科学，2011，38(12)：167-171．(Ouyang Chunping, Hu Changjun, Li Yang, et al. Approach of Ontology Learning from Relational Database Based on FCA[J]. Computer Science, 2011, 38(12): 167-171.)
[10] 滕广青，毕强．基于概念格的异构资源领域本体构建研究[J]．现代图书情报技术，2011(5): 7-12. (Teng Guangqing, Bi Qiang. A Study on Domain Ontology Construction from Heterogeneous Resources Based on Concept Lattice [J]. New Technology of Library and Information Service, 2011(5): 7-12.)
[11] 黄奇，郑建国．基于语义分析的产品分类本体查询研究[J]．情报学报，2015, 34(4): 398-413. (Huang Qi, Zheng Jianguo. Ontology Query for Product Classification Based on Semantic Analysis[J]. Journal of the China Society for Scientific and Technical Information, 2015, 34(4): 398-413.)
[12] Noy N F, McGuinness D L. Ontology Development 101: A Guide to Creating Your First Ontology [R]. Stanford: Stanford Knowledge Systems Laboratory Technical Report, 2001: 1-25.
[13] Uschold M, Gruninger M. Ontologies: Principles, Methods and Applications [J]. Knowledge Engineering Review, 1996, 11(2): 93-136.
[14] National Cancer Institute. OWLFCAView Manual [EB/OL]. [2015-05-02]. https://wiki.nci.nih.gov/display/Protege/OWLFCAView+Protege+Tab+Plug-in+Manual.
[15] eCl@ss. eCl@ss Wiki [EB/OL]. [2015-06-08]. http://wiki.eclass.eu/wiki/Main_Page.
[16] Gu T, Wang X H, Pung H K, et al. An Ontology-Based Context Model in Intelligent Environments [C]. In: Proceedings of Communication Networks and Distributed Systems Modeling and Simulation Conference. 2004: 270-275.

作者贡献声明:

陆佳莹，袁勤俭，黄奇，钱韵洁：提出研究思路，设计研究方案，论文最终版本修订；
陆佳莹：进行实验，采集、处理和分析数据，起草论文。

利益冲突声明:

所有作者声明不存在利益冲突关系。

支撑数据:

支撑数据见期刊网络版 http://www.infotech.ac.cn。
[1] 陆佳莹，袁勤俭，黄奇，钱韵洁．eclass_medical.rdf-xml.owl. 实例中的领域本体 owl 文档.
[2] 陆佳莹，袁勤俭，黄奇，钱韵洁．eclass_class_adjustment.doc. 34 大类类目调整具体方案.
[3] 陆佳莹，袁勤俭，黄奇，钱韵洁. eclass_34_ch-en.csv. 34 大类中英文对照表.

收稿日期: 2015-12-27
收修改稿日期: 2016-02-28

Building Product Domain Ontology with Concept Lattice Theory

Lu Jiaying[1] Yuan Qinjian[1] Huang Qi[2] Qian Yunjie[1]
[1](School of Information Management, Nanjing University, Nanjing 210093, China)
[2](State Information Resources Management Research Base, Nanjing University, Nanjing 210093, China)

Abstract: [**Objective**] The paper proposes a method to create new Ontology for reusable products based on the concept lattice theory, aiming to improve the traditional methodology. [**Methods**] First, we extracted classes, attributes and relationships data from product categories and made appropriate adjustments. Second, we created formal contexts and concept lattices to generate product Ontology automatically. [**Results**] The proposed method improved the product classification scheme and enhanced the reusability and sharing of Ontology. It also visualized the existing and potential entities as well as their relationships. [**Limitations**] Only used a segment of medical product Ontology to implement the proposed method, and then discussed its flexibility for eCl@ss product classification system. [**Conclusions**] The field and scope of Ontology need to be clearly defined before we construct product Ontology based on concept lattice. The categories should be sorted and grouped with scientific classification principles. We also need to build the property dictionary and define object properties. The Ontology will be generated automatically with formal context and concept lattice.

Keywords: Concept Lattice Domain Ontology Ontology construction Product Ontology

Plum Analytics 引入临床引用

EBSCO 子公司 Plum Analytics 一直以来持续不断的探索出测量研究影响力的方法，近日，其为临床影响引入了一个新的影响力指标：临床引用。当研究成果以临床目的被引用时，临床和转化科学领域的科研人员，以及那些资助和出版他们成果的工作人员，都将能够利用 PlumX 工具组件跟踪他们研究的影响情况。

PlumX 工具组件是由 5 个 Altmetrics 产品所组成的一个系列工具组件，它能够提供评价和追踪研究的方法，并进而发现未来的受资助机会。这些产品追踪的指标包括以下五大类：使用(Usage)、获取(Captures)、提及(Mentions)、社交传媒(Social Media)和引用(Citations)。引用类别传统上只包含了一些学术指标。Plum Analytics 是第一个将指标扩大到临床引用上的 Altmetrics 公司，临床引用是指临床信息资源中的引用情况，可以表征临床影响。

Plum Analytics 的创始人和主席 Andrea Michalek 认为临床引用为这些临床研究和转化科学研究人员、资助方和出版社工作人员提供一个额外的方式来评价他们的研究。“Plum Analytics 的目的是建造新的方式来评价和谈论研究及其影响。临床影响是研究成果影响力的一个非常重要的方面，其常常被忽视或低估。临床引用是 Plum Analytics 在评估影响力方面的一个最新发明。临床引用能帮助那些进行临床研究和转化科学研究的研究人员及时的发布他们的研究故事，并且为他们的研究获得相应的荣誉。”

PlumX 中的第一个临床引用数据来自 DynaMed Plus，其他的数据库也将逐步采用。临床医生依赖 DynaMed Plus 提供对于最新实证信息的即时访问，以保证对于病人做出最合理的医疗决定。PlumX 对 DynaMed Plus 中的主题进行了文本挖掘，并监控这些主题引用特定研究成果的时间。如此发现一条引用，PlumX 则为该研究计算一次临床引用。如果多个主题引用了相同的研究，则该研究将获得更高的临床影响计数。研究人员能在 PlumX 仪表盘中的 DynaMed Plus 中看到一系列的主题及其链接。

(编译自：https://www.ebscohost.com/careers/article/clinical-citations-from-plum-analytics-do-what-traditional-citations-cannot)

(本刊讯)